**Requested Review of Procedures**
**for the MAWQ/UMD Best Management Practice Project**

**STAC BMP Efficiencies Task Group**

Jim Pease
Mary Beth Adams
Saied Mostaghimi
Mark Walbridge
Dave Hansen

July 19, 2007

The Task Group was requested to support the Mid-Atlantic Water Quality Program/University of Maryland (MAWQ/UMD) Best Management Practice Project by providing review and assessment of the process whereby MAWQ/UMD arrived at Best Management Practice (BMP) efficiency recommendations. Specifically, the June 19, 2007 letter from MAWQ/UMD (copy attached), requests: 1) review of "..the relative efficiency of the BMPs across sectors.." and 2) review of "..the logic and process that was used to develop BMP definitions and efficiencies."

Our interpretation of this charge should be clearly stated. We are not requested to nor will we state judgments about the magnitude of MAWQ/UMD recommended BMP efficiencies. Thus, we can state no judgments concerning the relative magnitudes of BMP efficiencies across sectors. The absolute magnitude of each BMP efficiency reduction is assumed to reflect the reduction in Total Nitrogen (TN), Total Phosphorus (TP), and Total Suspended Solids (TSS) on a land surface achieved through implementation of a BMP compared to that land surface without the BMP. The relative ranking of BMP effectiveness (and the implication of adjustments if those of some sectors seem "unreasonable") is not a scientific issue. The issue is scientific confidence in the absolute magnitude of each BMP efficiency reduction, which leads us to the second task, that of reviewing the logic and process whereby MAWQ/UMD assessed recommendations made by experts and, in some cases, modified such recommendations.

We have reviewed and discussed the document entitled, "Developing BMP Efficiencies for Tributary Strategies" (attached). We paid particular attention to the criteria that were used to assess and to justify, in some cases, changes in current and expert-recommended BMP efficiencies. Although the criteria are not collected in the document, we have focused on several statements (emphasis added):

Criterion 1. "We are recommending efficiencies based on **operational conditions**."
Criterion 2. "..does the recommended efficiency use **negative efficiencies** in its calculations?"
Criterion 3. "..some overview and adjustment of all recommendations must occur to be **consistent** among BMPs."

Criterion 4. "**Peer reviewed literature was given more weight** than design standards/manuals although both were considered in BMP development."

Criterion 5. "**Single site studies should be utilized over multi site studies**."

Other comments that were relevant to our task include:

1. "**Best professional judgment** was used to discount efficiencies to reflect the variability of operational systems.
2. BMP efficiencies "..**do not consider the variability and uncertainty associated with rate of implementation, operation and maintenance, replacement, spatial and management variability or tracking and reporting.**

The Task Group was unclear concerning the issues of Criterion 5 (Single site). There are studies that examine one BMP at one site, one BMP at several sites, multiple BMPs at one site, or multiple BMPs at several sites. Obviously, more confidence can be expressed about studies in which all factors are controlled except one (a specific BMP). Conducting such a controlled analysis on multiple sites provides more information on the spatial variability of BMP effectiveness, which we regard as valuable. The effects of a particular BMP within an experiment utilizing multiple BMPs is much more difficult to ascertain, and more so the single-BMP effect on multiple sites.

Criterion 3 (Consistency) refers to the charge given to the experts, and to their professional diligence and thoroughness. MAWQ/UMD states that some experts used the lack of research data to justify deep discounts of the few reported efficiencies, while other experts refused to change current efficiencies because of the lack of research data. We would hope that such a situation was anticipated, and that the charge to the expert specifically stated how such situations were to be handled. If experts ignored such instructions, then they have not performed well in a professional manner.

We had few specific comments on Criteria 1, 2, and 4, which also serve well as principles. The Chesapeake Bay model must be calibrated to function with operational rather than research BMP efficiencies. Hence, if reported negative efficiencies reflect operational conditions, they should be considered in an assessment of the BMP efficiency literature. Peer-reviewed literature has more credibility than do design standards/manuals which have not been subjected to independent examination.

The Task Group considers that more attention should have been paid to an assessment of the research quality of reported BMP efficiencies. How representative were the research site characteristics to conditions in the Chesapeake Bay watershed? Are their objective justifications for using a single BMP efficiency across a wide spatial area? What is the duration of the experiment, and how critical is the duration to the reported efficiency results? Were reported results arrayed and the median value chosen to reduce the effects of variability and uncertainty? Do results reflect changes in reduction efficiencies over the lifetime of the BMP? Such as assessment would be particularly useful with regard to

Forest Harvesting Practices, which relies extensively on one Coastal Plain study, and is not likely to represent well the efficiencies obtainable throughout the watershed.

The Task Group considers that the key missing element in the MAWQ/UMD process is **justification of the magnitude of adjustments.** The Task Group is not privy to the process whereby, for example, the expert recommendations for TN, TP, and TSS for Off-stream Watering With Fencing were adjusted to the specific numbers 25, 40, and 40 percent, respectively. Why not 30, 45, and 45, or 20, 30, and 30? It is stated that the expert recommendations were not consistent with other efficiencies of similar agricultural BMPs, but justification for these specific reduction percentages is not given. We do not have sufficient information to judge the Best Professional Judgment cited as the process used for selection of these particular efficiencies.

MARYLAND

College Park, MD 20742
Phone 301-405-1306
FAX 301-314-2763

*Department of Environmental Science and Technology*

June 19, 2007

Dear STAC review team,

Thank you for agreeing to continue your support of the Mid-Atlantic Water Quality Program – University of Maryland (MAWQ-UMD) Best Management Practice (BMP) project. At last week's STAC meeting, you agreed to carry out the following two tasks by July 27, 2007:

**Task 1: Review the MAWQ-UMD relative rankings of BMPs.**
The first task is for the task force within STAC to review the entire suite of MAWQ-UMD BMP recommended efficiencies to ensure proposed efficiencies make sense across sectors (i.e., urban, agriculture, etc.) and are appropriately ranked. Please review the list of BMPs that are ranked according to their effectiveness of reducing nitrogen, phosphorus, and sediment across all sectors. We believe it is important to assess the relative efficiency of the BMPs across sectors as a "cross check" to ensure that the results of this project make intuitive sense and imply a consistent thought process across sectors, based on what is known about the performance of these BMPs.

**Task 2: Review MAWQ-UMD Process for Making BMP Recommendations**
Later this week I will send you a detailed description of the logic and process the MAWQ-UMD used to develop BMP definitions and efficiencies. As you review our logic, we ask that you consider the following questions for specific BMPs that are under debate in the Nutrient Subcommittee source workgroups:

Urban Wetlands and Wetponds: The Urban Stormwater Workgroup is likely to propose higher efficiencies based on only the multi-studies, ruling out the individual studies that had negative efficiencies. Please carefully evaluate the MAWP-UMD logic for including the studies with negative efficiencies and give us an indication on whether you think these studies should be considered or not.

Forest Buffers: The Forestry Workgroup is likely to propose higher efficiencies (such as 80%) which are supported by a number of studies. The MAWP-UMD recommendations are lower because we think that operationally across the watershed we will not see these high reductions. Initial model runs support using lower efficiencies. Please evaluate our logic in making this decision.

Please keep your comments limited to the scope of work and funding for the project. Time is a limiting factor; we have to provide recommendations on efficiencies for calibration of the watershed model by mid-August. In addition, there is limited money in the grant and we could not fund every model run we wished to evaluate. Ultimately we were instructed to review the literature and model runs in order to serve as the interpreter between the experts, reviewers and Chesapeake Bay Program to ensure a consistent product. This is necessary because the expert who developed the BMP reports may not understand the constraints of the model and different developers had different approaches when recommending efficiencies. Finally, we are tasked with recommending efficiencies that are reflected of operational conditions and can be applied across the whole watershed.

**Attachments:**

1. List of ranked BMP efficiencies from the MAWQP-UMD.
2. Technical documentation for each BMP.

Attached are the full detailed reports for each BMP including the expert recommendations, the STAC review, and the MAWP-UMD final recommendation. Please note that we have not included cover crops at this time, since we are still in the process of developing efficiencies. Also, we have decided to postpone the urban stream restoration and rotational grazing BMP un year 2 of the project.

Please visit http://www.mawaterquality.org/temporary/bmp.htm to download the files.

Again thank you for your time and energy. Your input is a significant piece of this project and will be used by the Tributary Strategy Workgroup and Nutrient Subcommittee during their review in August. Please don't hesitate to contact me or Tom with any questions.

## Developing BMP Efficiencies for Tributary Strategies

### Background

Uncertainty in nonpoint source estimates is due to several factors including but not limited to variability in precipitation, hydrology and geology; variable performance of land management practices; lag time between implementation of practices and full performance and between implementation and observed water quality benefits; and the effect of cover and slope on pollutant load delivery to receiving waters. To avoid the errors and shortcomings in BMP efficiency estimation, and to more realistically estimate real world pollutant removals from BMPs, one must examine this suite of factors. These factors should be used to adjust efficiencies estimated from research plots. Not every BMP will be subject to all the factors, but a research project will not capture the suite of factors that determine efficiencies in natural systems.

The expected spatial variability for a practice should be estimated based on available science and knowledge of the expected geographic extent of implementation of the practice. Different reduction efficiencies should be established for practice implementation across different physiographic, geomorphic or hydrologic settings. Where possible, efficiencies should be adjusted for surface water and groundwater interactions (permeability), along with geology and soil types (slope, seeps, floodplain, etc.).

The loss pathways and hydrologic lag time associated with each practice should be examined to see if an adjustment in effectiveness should be made. Transport of pollutants occurs through a variety of environmental pathways that include the soil surface, vadose zone, saturated zone, tile drains, and streams. The time scale of this transport varies substantially depending on the pathway followed by water from the land surface to the stream. For example, surface runoff to a stream may take minutes to days, whereas leaching to groundwater followed by discharge to a stream may take months to decades.

As the STAC paper (Simpson et al., 2004) reported, hydrologic lag times are well documented in the Chesapeake Bay watershed. Dissolved nitrogen associated with groundwater may have a transport time of years to decades, with a median time of about 10 years (Lindsey et al., 2003). Nutrients associated with sediment can have much longer transport times (several decades) in the watershed because of their storage in soil and stream corridors, both of which are greatly influenced by annual rainfall. Additionally, the location of the source area in the watershed will influence the lag time between implementation lag times and improvement in bay water quality. Planning, implementation, and practice maturity lag times may be easier to estimate than hydrologic lag times, but are rarely considered.

BMP efficiencies should match the practice implementation schedule. Many practices are reported as implemented once the plan or design has been completed. In reality, the plan may call for phased implementation over as much as five to ten years. In addition, with agricultural land the farmer may not implement the practice as scheduled due to climatic, management or economic constraints. The time it takes for an implemented practice to reach its full potential may delay pollution reduction percentages. Efforts should be made to assure that reported implementation is close to actual, and to determine if implementation and operation is as rigorous as specified in the practice. Identifying possible lag times in reaching BMP pollution reductions due to phased-in implementation or time to maturity will accurately estimate effectiveness.

Efficiencies will change from the research/demonstration scale to the watershed/basin scale. Both the scale difference and the management differences between a research plot and a BMP site will alter efficiencies. On a research site, the BMP is designed, operated and maintained in a very controlled manner. This ensures that the BMP is achieving its full potential or near its highest efficiency. On a watershed scale, the same level of control and oversight is not possible, thus not ensuring the BMP is fully functioning.

The nature of plot, field and watershed scale research introduces variability in BMP effectiveness. As mentioned at the plot scale, the researcher controls the land. Typically only one experiment is carried out at a time. Varying levels of treatment, including controls, are easily carried out in a replicated experimental design. Research designs include approaches that reduce the likelihood of inconclusive results due to variations in natural factors such as soil, hydrology, topography, and other conditions. Most aspects of weather are consistent from plot to plot, and rainfall is often simulated, providing control over amount and intensity. Data is analyzed statistically to account for variability and significance of results.

At the field scale, research becomes more difficult as replication becomes less feasible or more expensive. Different levels of treatment are still feasible and each field receives a uniform treatment across its full extent. Heterogeneity in soils, topography, weather, and management introduce larger errors into the observations, obscuring the effects of the treatments to a greater extent than the plot scale. Rainfall is not simulated, and is natural, resulting in heterogeneous amounts and intensity across the research site.

At the watershed scale, the researcher becomes more of an observer than a manipulator of the research site. Most water quality research projects attempt to interpret the cumulative result of multiple changes in land management practices taking place at different times. Replication of experiments is rarely feasible. Implementation of specific practices usually cannot be targeted to specific places in the landscape, and is often limited to a small percentage of the total land area. Timing and intensity of climatic events is often the main determinant of fluctuations in water quality. Weather and the agricultural economy play a large role in crop choices, tillage practices, and fertilizer application. If a control watershed is available, the researcher often has little control over management. There may be lag times between land use change and a response in water quality. Given the high level of natural variability in water quality data, failure to detect a change is not an indication that BMPs did not work. Alternatively, given the multitude of factors that influence water quality, detecting a change does not lead to the conclusion that the BMPs were responsible for the change unless other factors, such as management changes, can be ruled out. All these problems become more severe as watershed size increases. The scale of study should be taken into account and be reflected in efficiency adjustment as research and demonstration site derived efficiencies for watershed scale implementation do not reflect the spatial viability of the entire watershed.

Extreme climatic events can have major impacts on BMP function and efficiency in events above its designed maximum. When data is available, the practice efficiency should be adjusted for events approaching, but within, the design maximum. Also, different lengths or widths of the BMP (where applicable, for example Riparian Forest Buffers) will alter efficiencies.

Watershed management conditions, including operation and maintenance of BMP, construction supervision, and/or upland land use change will also impact efficiencies, usually making them lower than research scales. While there is little quantitative information on how BMP efficiencies should be adjusted to account for the impacts of improper maintenance on receiving waters, general adverse impacts on practice operation are understood. If maintenance is neglected a BMP may become impaired, no longer providing its designed functions. Proper maintenance of outlet structures, flow splitters and clean out gates is key to achieving a stormwater BMPs designed efficiency (Koon, 1995).

In addition, sediment accumulation is one maintenance concern that if not addressed may adversely affect the effectiveness of some BMPs. As sediment accumulates it decreases storage volume and detention time, bypassing the intended functions of the BMP and increasing discharge of nutrient and sediment rich stormwater (Livingston et al., 1997). Increased discharge will lead to decreased downstream channel stability, resulting in increased sediment loads and a reduction in available aquatic habitat. The consequences of increased stormwater discharges from sediment filled BMPs, are a reduction in the BMPs pollution removal efficiencies, and ultimately, increased ecological impairments. The uncertainty in how improper maintenance will adjust BMP efficiencies supports the recommendation to use a more conservative percent removal estimate.

Model output and monitoring data must be consistent and used appropriately. Better research on demonstration and monitoring of BMP, system and small watershed conservation effects will increase confidence in BMP effectiveness. Finally, managers, policy makers, and involved citizens must be made aware of potential implications of adaptive science and understand why an adaptive approach is essential.

As discussed, BMP efficiency estimation is not an exact science. Experience has shown that most water quality BMPs do not perform as well with widespread implementation as projected based on research. As such, any program using BMP efficiencies as an estimate of pollution discharges should reduce efficiencies to build uncertainty into its BMP reduction estimates. TMDLs and trading permits are designed to result in pollution discharges that meet legal water quality concentrations. If they use BMP efficiency estimates to achieve permit or TMDL requirements, then efficiency estimates should be adequately conservative to assure needed water quality conditions are met.

**Factors Considered in Efficiency Development**

When analyzing BMP effectiveness an adaptive management approach is warranted. Methodology employed to develop efficiencies varies depending on the practice, but there are guidelines the Mid-Atlantic Water Quality Program (MAWQ) utilized when recommending efficiencies. These criteria help determine how efficiencies reported in literature should be adjusted to reflect operational conditions.

As discussed, data at any scale is limited as research, field and watershed scale efficiencies differ. Thus assuming all BMPs will produce the same efficiency at the operational scale as the research scale is erroneous. We are recommending efficiencies based on operational conditions. Is the recommendation by the expert and/or in literature representative of the efficiency one would expect at the watershed scale? Does this efficiency represent watershed-wide effectiveness? If the efficiency is not expected to occur uniformly across the watershed then a more conservative efficiency was assigned. Best professional judgment is used to adjust expert efficiencies because the scientific review usually only considered research values. These values, however, are not reflective of operational conditions.

Some studies report negative efficiencies due to natural or construction and operational related issues with the BMP. During the MAWQ/UMD efficiency development process, negative efficiencies reported in literature are not omitted because they occur in real world situations. The question then becomes, does the recommended efficiency use negative efficiencies in its calculations? Are studies with negative efficiencies used to calculate medians or eliminated from the study sample? If negative efficiencies are not included, efficiencies should be discounted to account for failed systems. Negative efficiencies often are not published and when they are, they have undergone rigorous scientific review.

As multiple experts are recommending efficiencies for their BMPs in their area of expertise, naturally the approach by each reviewer varies. Thus some overview and adjustment of all recommendations must occur to be consistent among BMPs. Ranking exercises will highlight inconsistencies used when various experts recommend efficiencies and adjustments are made accordingly. Some experts used the lack of data to justify not changing current efficiencies, while others used the lack of data to justify significantly reducing efficiencies.

For example, one expert was very critical during his review of off-stream watering BMPs and he justified reducing literature-based efficiencies by 50% due to the lack of data. His recommendations lowered the effectiveness to such a degree that when compared to other BMPs during a ranking exercise the results did not intuitively make sense. Thus his recommendations were deemed too severe and not used. Another reviewer was hesitant to make any recommendations without sufficient literature, stating a change to the current efficiencies could not be justified due to the lack of data. MAWQ used his review to develop recommendations. The developer strictly evaluated the efficiencies from a scientific standpoint and stated he made no adjustment for factors that discount efficiencies (operational versus research differences). As such efficiencies were higher than would be expected to be observed operationally.

Some current BMP efficiencies were developed with limited data. In situations where best professional judgment had to be used, any new literature was considered in refining the efficiency. Reviews of the studies used, however, were critical in these cases. Alternatively, for BMPs that had sufficient/adequate data used to develop current efficiencies, UMD/MAWQ required a large body of consistent data to motivate a refinement to BMP.

Peer reviewed literature was given more weight than design standards/manuals although both were considered in BMP development. Peer reviewed literature has undergone a robust critical screening before it is published. Non peer reviewed literature is not submitted to the same screening process.

Single site studies should be utilized over multi site studies, as the former study of individual BMP project sites, while the later uses design ratings for particular BMPs based on multiple sites or professional judgment. These multi-site studies may incorporate the efficiencies reported in the single site studies thus counting some studies twice during statistical calculations.

Site specific conditions will create variability in efficiencies:

- Soils
- Hydrology – surface and subsurface flow
- Size (width, length)
- Watershed management conditions
- Specie composition
- BMP age and time to maturity
- Climate and other seasonal changes

Unfortunately specific data that describes how efficiencies should be adjusted for the aforementioned criteria is limited. Best professional judgment was used to discount efficiencies to reflect the variability of operational systems.

It is important to note that these criteria do not consider the variability and uncertainty associated with rate of implementation, operation and maintenance, replacement, spatial and management variability or tracking and reporting. These items that adjust efficiencies need to be investigated and applied to future efficiency refinement procedures.

**Specific Workgroup Concerns**

From the presentations and discussions so far it is fair to assume that the workgroups will recommend different efficiencies than the UMD/MAWQ Program numbers. Some examples of the issues are:

USWG suggested that studies that report negative efficiencies be eliminated from the analysis. They stated that negative efficiencies represent a failed system and should not be included. These negative efficiencies could be the result of errors in design, construction, operation and maintenance or are within the expected range of function of the BMP. UMD-MAWQ feels these situations do occur operationally and so negative efficiencies should be included.

One specific example is the urban wetlands and wet ponds report. The expert developer for this BMP conducted a literature synthesis and broke the statistical analysis out by single site studies (study of individual BMP project sites) and muti-site studies (design ratings for particular BMPs based on multiple sites or professional judgment). The developer and the STAC reviewer stated the values closer to the mean and median efficiencies of the single-site studies should be used to determine effectiveness than those of the multi-site studies. The USWG, however, wants to use the multi-site studies and not the single site ones to recommend efficiencies.

Multi-site analyses were used in the other urban BMP reports but preference was always given to the single site studies. The USWG wants to include the studies used to develop efficiencies from their design manuals in the MAWQ-UMD project. BMP projects from the Center for Watershed Protection database was used to develop VA draft regulations and MD and PA stormwater design manuals. These studies fall into the multi-site category and may not reflect operational conditions. Upon further evaluation of all sources considered in development of the urban wetland and wet pond practices, it was found that the developer had included the sources from the design manuals in his multi-site analyses. The analysis by the developer includes the median values for all 145 studies used in the 2000 version of the Center for Watershed Protection database. In addition, some single site studies from the database are also included in the developers single site analyses.

**Conclusion**

Developing efficiencies that reflect operational, real-world conditions requires a holistic view point. There are certain qualities of research studies that do not incorporate all the factors that will influence operational efficiencies. To account for this, research based efficiencies must be adjusted using the aforementioned guidelines.

**STAC Participation**

1) STAC agreed to review the relative ranking of the MAWQ/UMD's recommended efficiencies. A spreadsheet that lists the recommendations broken out by pollutant is available. There are a couple options listed for riparian buffers and the numbers that correspond to each option are provided in two separate spreadsheets, one for grass and one for forest buffers. The reports on each BMP are also provided.

2) STAC agreed to also review the MAWQ/UMD process for developing its recommendations. This paper is provided to aid in this evaluation.

**Reference:**

Koon, J. 1995. Evaluation of Water Quality Ponds and Swales in the Issaquah/East Lake Sammamish Basins. King County Surface Water Management Division, Seattle, WA. IN: Shoemaker, L., Lahlou, M., Doll, A., and P. Cazenas. 2002. Stormwater Best Management Practices in an Ultra-Urban Setting: Selection and Monitoring; Fact Sheet Detention Ponds. Federal Highway Administration, Landover, Maryland. http://www.fhwa.dot.gov/environment/ultraurb/index.htm

Lindsey, B.D, Phillips, S.W., Donnelly, C.A., Speiran, G.K., Plummer, L.N., Bohlke, J.K., Focazio, M.J., Burton, W.C., and Busenberg, E., 2003. Residence Times and Nitrate Transport in Ground Water Discharging to Streams in the Chesapeake Bay Watershed. Water-Resources Investigations Report 03-4035, 201p.

Livingston, E.H., Shaver, E., Skupien, J.J., and R.R. Horner. 1997. Operation, Maintenance, and Management of Stormwater Management Systems. Watershed Management Institute, Inc., Ingleside, MD.

Simpson, T.W., Musgrove, C.A., and Korcak, R.F, 2004. Innovation in Agricultural Conservation for the Chesapeake Bay: Evaluation Progress and Assessing Future Challenges. The Scientific and Technical Advisory Committee, Chesapeake Bay Program, Annapolis, MD. Available Online <http://www.chesapeake.org/stac/stacpubs.html>